PCT/EP 00 / 02701

Europäisches Patentamt

Eur pean Patent Office

Office eur péen des brevets

REC'D 03 JUL 2000

WIPO PCT

EPO - DG 1

02. 06. 2000

(70)

Bescheinigung Certificate Attestation

4

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

99111321.8

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

Hatten

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN
THE HAGUE,
LA HAYE, LE

12/04/00

EPA/EPO/OEB Form 1014 - 02.91

Europäisches Patentamt | European Patent Office | Office européen des brevets

Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.: 99111321.8
Demande n°:

Anmeldetag:
Date of filing: 10/06/99
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):

Stymne, Sten
268 90 Svalöv
SWEDEN
Dahlqvist, Anders
244 66 Furulund
SWEDEN

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:

A new class of enzymes in the biosynthetic pathway for the production of triacylglycerol and recombinant DNA molecules encoding these enzymes

In Anspruch genommene Prioriät(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat: State: Pays:	EP	Tag: Date: Date:	01/04/99	Aktenzeichen: File no. Numéro de dépôt:	EPA 99106656

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:

/

Am Anmeldetag benannte Vertragstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE
Etats contractants désignés lors du depôt:

Bemerkungen:
Remarks:
Remarques:

Further applicants are : Ulf STAHL, 753 24 UPPSALA, Sweden
Marit LENMAN, 233 59 LUND, Sweden
Antoni BANAS, 08-110 SIEDLCE, Poland
Hans RONNE, 756 54 UPPSALA, Sweden

A new class of enzymes in the biosynthetic pathway for the production of triacylglycerol and recombinant DNA molecules encoding these enzymes

Applicants and inventors:

Anders Dahlqvist, Ulf Ståhl, Marit Lenman, Antoni Banas, Hans Ronne and Sten Stymne

Description

The present invention relates to the isolation, identification and characterization of recombinant DNA molecules encoding enzymes catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol.

The invention further relates to novel type of enzymes, the genes encoding such enzymes and the use of these genes for transformation. More specifically, the invention relates to use of a previously un-described type of enzymes hereinafter designated phospholipid:diacylglycerol acyltransferases (PDAT) and the genes encoding them. This type of genes expressed alone in transgenic organisms will enhance the total amount of oil (triacylglycerols) produced in the cells. The PDAT genes, in combination with a gene for the synthesis of an uncommon fatty acid will, when expressed in transgenic organisms, enhance the levels of the uncommon fatty acids in the triacylglycerols.

There is considerable interest world-wide in producing chemical feedstock, such as fatty acids, for industrial use from renewable plant resources rather than non-renewable petrochemicals. This concept has broad appeal to manufacturers and consumers on the basis of resource conservation and provides significant opportunity to develop new industrial crops for agriculture.

There is a diverse array of unusual fatty acids in oils from wild plant species and these have been well characterised (see e.g. Badami & Patil, 1981). Many of these acids have industrial potential and this has led to interest in domesticating relevant plant species to enable agricultural production of particular fatty acids.

Development in genetic engineering technologies combined with greater understanding of the biosynthesis of unusual fatty acids now makes it possible to transfer genes coding for key enzymes involved in the synthesis of a particular fatty acid from a wild species into domesticated oilseed crops. In this way individual fatty acids can be produced in high purity and quantities at moderate costs.

In all crops like rape, sunflower, oilpalm etc., the oil (i.e. triacylglycerols) is the most valuable product of the seeds or fruits and other compounds like starch, protein, and fibre is regarded as by-products with less value. Enhancing the quantity of oil per weight basis at the expense of other compounds in oil crops would therefore increase the value of crop. If genes, regulating the allocation of reduced carbon into the production of oil can be up-regulated, the cells will accumulate more oil on the expense of other products. Such genes might not only be used in already high oil producing cells such as oil crops but could also induce significant oil production in moderate or low oil containing crops such as e.g. soy, oat, maize, potato, sugarbeats, and turnips as well as in micro-organisms.

Summary of the invention

Many of the uncommon fatty acids of interest, e.g. medium chain fatty acids, hydroxy fatty acids, epoxy fatty acids and acetylenic fatty acids, have physical properties that are distinctly different from the common plant fatty acids. The present inventors have found that, in plant species naturally accumulating these uncommon fatty acids in their seed oil (i.e. triacylglycerol), these acids are absent, or present in very low amounts in the membrane (phospho)lipids of the seed. The low concentration of these acids in the membrane lipids is most likely a prerequisite for proper membrane function and thereby for proper cell functions. One aspect of the invention is that seeds of transgenic crops can be made to accumulate high amounts of uncommon fatty acids if these fatty acids are efficiently removed from the membrane lipids and channelled into seed triacylglycerols.

The inventors have identified a novel class of enzymes in plants catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the production of triacylglycerol and, presumably, lysophospholipids and that these enzymes (phospholipid:diacylglycerol acyltransferases abbreviated as PDAT) are involved in the removal of hydroxylated, epoxygenated fatty acids, and probably also other uncommon fatty acids such as medium chain fatty acids, from phospholipids in plants. Further, the same enzyme reaction was shown to be present in microsomal preparations from baker's yeast (*Saccharomyces cerevisiae*). A so called ‚knock out‘ yeast mutant, disrupted in the respective gene was obtained and microsomal membranes from the mutant was shown to totally lack PDAT activity. Thus, it was proved that the disrupted gene encodes a PDAT enzyme (SEQ ID NO: 1 and 2). In addition, a gene from *Schizosaccharomyces pombe*, SPBC776.14 (SEQ ID. NO. 3), having 46% identity over 571 amino acids, and three *Arabidopsis thaliana* genomic sequences coding for putative proteins with amino acid sequence having approximately 44 % identity over a total of 378 amino acids (SEQ ID NO: 4), 47 % identity over 73 amino acids (SEQ ID NO: 9), and 33% identity over 105 amino acids (SEQ ID NO. 10), with the yeast enzyme, were

identified.

Also, a partially sequenced cDNA clone from *Neurospora crassa* (SEQ ID NO. 8) showed 42% identity over 167 amino acids, and a *Zea mays* est clone (SEQ ID NO. 6 and 7) showed 36% identity over 96 amino acids, with the yeast enzyme were identified. Finally, two cDNA clones were identified, one *Arabidopsis thaliana* est (SEQ ID NO. 5) which was 98% identical in nucleotide sequence to the *Arabidopsis thaliana* genomic sequence (ID no. 4), and a *Lycopersicon esculentum* est clone (SEQ ID NO. 11) showed 84% identity in 72 nucleotides to Seq ID No. 10.

In a first embodiment, this invention is directed to a PDAT enzyme and nucleic acid sequences that encode a PDAT. This includes biologically active PDATs and encoding sequences for PDATs as well as sequences that are to be used as probes, vectors for transformation or cloning intermediates. The PDAT encoding sequence may encode a complete or partial sequence depending upon the intended use. All or a portion of the genomic sequence, cDNA sequence, precursor PDAT or mature PDAT is intended.

In a different aspect, this invention relates to a method for producing a PDAT in a host cell or progeny thereof, including genetically engineered oil seeds, yeast and moulds or any other oil accumulating organism, via the expression of a construct in the cell. Cells containing a PDAT as a result of the production of the PDAT encoding sequence are also contemplated within the scope of the invention.

In a different embodiment, this invention also relates to methods of using a DNA sequence encoding a PDAT for increasing the oil-content within a cell.

Another aspect of the invention relates to the accommodation of high amounts of uncommon fatty acids in the triacylglycerol produced within a cell, by introducing a DNA sequence producing a PDAT that specifically removes these fatty acids from the membrane lipids of the cell and channel them into triacylglycerol. Plant cells having such a modification are also contemplated herein.

A PDAT of this invention includes any sequence of amino acids, such as a protein, polypeptide or peptide fragment obtainable from a microorganism, animal or plant source that demonstrates the ability to catalyse the production of triacylglycerol from a phospholipid and diacylglycerol under enzyme reactive conditions. By,enzyme reactive conditions" is meant that any necessary conditions are available in an environment (e.g., such factors as temperature, pH, lack of inhibiting substances) which will permit the enzyme to function.

Other PDATs are obtainable from the specific sequences provided herein. Furthermore, it will be apparent that one can obtain natural and synthetic PDATs, including modified amino acid sequences and starting materials for synthetic-protein modelling from the exemplified PDATs and from PDATs which are obtained through the use of such exemplified sequences. Modified amino acid sequences include sequences that have been mutated, truncated, increased and the like, whether such sequences were partially or wholly synthesised. Sequences that are actually purified from plant preparations or are identical or encode identical proteins thereto, regardless of the method used to obtain the protein or sequence, are equally considered naturally derived.
The nucleic acid sequences (DNA and RNA) of the present invention can be used to screen and recover "homologous" or ,"related" PDATs from a variety of plant and microbial sources.
Further, it is also apparent that a person skilled in the art can, with the information provided in this application, in any organism identify a PDAT activity, purify an enzyme with this activity and thereby identify a ,non-homologues" nucleic acid sequence encoding such an enzyme.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1.

PDAT activity in microsomal fractions of S. cerevisiae. Aliquots of lyophilised microsomal membranes (10 nmol phosphatidylcholine) from a wild type yeast (strain YN979) (lane 1-3, 3), a yeast mutant (strain B10280), disrupted in the YNR008w gene (lane 4-6, 9) or the yeast mutant complemented with a single copy plasmid containing the PDAT gene (lane 7) were assayed for PDAT activity. 2 nmol sn-1-oleoyl-sn-2-[^{14}C]-ricinoleoylphosphatidylcholine (lane 1-7) or sn-1-oleoyl-sn-2-[^{14}C]-oleoyl-phosphatidylcholine (lane 8-9) and 5 nmol of dioleoyl-diacylglycerol (lane 2,5, 7-9) or rac-oleoyl-vernoloyldiacylglycerol (lane 3, 6) were added in benzene solution. The benzene was evaporated under N_2 (g) and 0.1 ml of 50 mM potassium phosphate, pH 7.2, was added. The suspension was thoroughly mixed and after 90 min at 30 C the lipids were extracted in chloroform and separated an thin layer chromatography on silica gel 60 plates in hexan/dietylether/acetic acid (35:70:1.5). The radioactive lipids were visualised and quantified an the plates by electronic autoradiagraphy (Instant imager, Packard, US). Abbreviations used: triacylglycerol, TAG, FA, fatty acid (i.e. oleic acid); 1-OH-TAG, monoricinoleoyl-triacylglycerol; 1-OH-1-epTAG, monoricinoleoyl-monovernoleoyl-triacylglycerol nd OH-FA, ricinoleic acid.

Brief Description of the SEQ ID:

SEQ ID NO. 1: Genomic DNA sequence of the *Saccharomyces cerevisiae* PDAT gene, YNR008w, genebank nucleotide ID number 1302481, and the suggested YNR008w amino acid sequence.

SEQ ID NO. 2: The suggested amino acid sequence of the yeast gene YNR008w from *Saccharomyces cerevisiae*.

SEQ ID NO. 3: Genomic DNA sequence of the *Schizosaccharomyces pombe* gene SPBC776.14.

SEQ ID NO. 4: Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with genebank accession number AB006704.

SEQ ID NO. 5: Nucleotide sequence and the corresponding amino acid sequence of the *Arabidopsis thaliana* est clone with genebank accession number T04806, and ID number 315966.

SEQ ID NO. 6: Nucleotide and amino acid sequence of the *Zea mays* cDNA clone with genebank ID number g4388167.

SEQ ID NO. 7: Amino acid sequence of the *Zea mays* cDNA clone with genebank ID number g4388167.

SEQ ID NO. 8: DNA sequence of part of the *Neurospora crassa* cDNA clone W07G1, ID number g4241729.

SEQ ID NO. 9: Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with genebank accession number AC004557.

SEQ ID NO. 10: Genomic DNA sequence of part of the *Arabidopsis thaliana* locus with genebank accession number AC003027.

SEQ ID NO. 11: DNA sequencce of part of the *Lycopersicon esculentum* cDNA clone with genebank accession number AI486635.

The present invention can be essentially characterized by the following aspects:

1. Use of a PDAT PDAT gene (genomic clone or cDNA) for transformation.
2. Use of a DNA molecule according to item 1 wherein said DNA is used for transformation of any organism in order to be expressed in this organism and result in an active recombinant PDAT enzyme in order to increase oil content of the organism.
3. Use of a DNA mol cule of item 1 wherein said DNA is used for transformation of any organism in order to prevent the accumulation of undesirable fatty acids in the membrane lipids.

4. Use according to item 1, wherein said PDAT gene is used for transforming transgenic oil accumulating organisms engineered to produce any uncommon fatty acid which is harmful if present in high amounts in membrane lipids, such as medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.
5. Use according to item 1, wherein said PDAT gene is used for transforming organisms, and wherein said organisms are crossed with other oil accummulating organisms engineered to produce any uncommon fatty acid which is harmful if present in high amounts in membrane lipids, comprising medium chain fatty acids, hydroxylated fatty acids, epoxygenated fatty acids and acetylenic fatty acids.
6. Use according to item 1, wherein the enzyme encoded by said PDAT gene or cDNA is coding for a PDAT with distinct acyl specificity.
7. Use according to item 1 wherein said PDAT encoding gene or cDNA, is derived from *Saccharormyces cereviseae*, or contain nucleotide sequences coding for an amino acid sequence 30% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
8. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharormyces cereviseae*, or contain nucleotide sequences coding for an amino acid sequence 40% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
9. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharomyces cereviseae*, or contain nucleotide sequences coding for an amino acid sequence 60% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
10. Use according to item 1 wherein said PDAT encoding gene or cDNA is derived from *Saccharomyces cereviseae*, or contain nucleotide sequences coding for an amino acid sequence 80% or more identical to the amino acid sequence of PDAT as presented in SEQ. ID. NO. 2.
11. Use according to claim 1 wherein said PDAT encoding gene or cDNA is derived from plants or contain nucleotide sequences coding for an amino acid sequence 40% or more identical to the amino acid sequence encoded by the PDAT gene from *Arabidopsis thaliana* as presented in SEQ. ID. NO. 4, 9 or 10, or to the protein encoded by the fullength counterpart of the partial *Zea mays*, *Lycopersicon esculentum*, or *Neurospora crassa* cDNA clones.
12. Transgenic oil accumulating organisms comprising, in their genome, a PDAT gene transferred by recombinant DNA technology or somatic hybridization.
13. Transgenic oil accumulating organisms according to item 12 comprising, in their genome, a PDAT gene having specificity for substrates with particular uncommon fatty acid and the gene for said uncommon fatty acid.
14. Transgenic organisms according to item 12 or 13 which are selected from the group consisting of fungi, plants and animals.

15. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants.
16. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants and where said PDAT gene is expressed under the control of a storage organ specific promotor.
17. Transgenic organisms according to item 12 or 13 which are selected from the group of agricultural plants and where said PDAT gene is expressed under the control of a seed promotor.
18. Oils from organisms according to item 12 - 17.
19. A method for altering acyl specificity of a PDAT by alteration of the nucleotide sequence of a naturally occurring encoding gene and as a consequence of this alternation creating a gene encoding for an enzyme witb novel acyl specifity.
20. A protein encoded by a DNA molecule according to item 1 or a functional fragment thereof.
21. A protein of item 20 designated phospholipid:diacylglyceriol acyltransferase.
22. A protein of item 21 which has a distinct acyl specificity.
23. A protein of item 13 having the amino acid sequence as set forth in Seq id no. 2, (and the proteins encoded by the genes fullength or partial genes set forth in SEQ ID NO. 1, 3, 4, 5, 6, 7, 8, 9, 10 or 11) or an amino acid sequence with at least 30 % homology to said amino acid sequence.
24. A protein of item 23 isolated from Saccharomyces cereviseae.

EXAMPLE 1

Determination of substrate utilisation by the acyl-CoA independent acyltransferase in the synthesis of triacylglycerols in microsomal preparations of developing castor bean endosperm and bakers yeast.

Microsomal membranes prepared from developing endosperm of castor bean (Ricinus communis) catalyse the selective transfer of ricinoleoyl-(12-hydroxy-9-octadecenoyl) and vernoloyl(12-epoxy-9-octadecenoyl) groups from both diacylglycerols and phosphatidylcholine into triacylglycerols. The substrate utilisation was investigated in castor bean microsomes by using radioactive sn-1-oleoyl-sn-2-[^{14}C]ricinoleoyl-diacylglycerol (sn-2-[^{14}C]-ricinoleoyl-diacylglycerol) or sn-1-oleoyl-sn-2-(^{14}C)ricinoleoyl-phosphatidylcholine (sn-2-(^{14}C]-ricinoleoyl-phosphatidylcholine) together with different non-radioactive diacylglycerol (see Table 1).

The preparatian of microsomal fractions of developing castor bean endosperm and freeze drying of the microsomes were performed in known manner. Assays with addition of diacylglycerol and phosphatidylcholine substrates were performed. The results is presented in Table 1 and showed that if radioactive sn-2-[^{14}C]ricinoleoyl-diacylglyoerol was used as the only added

substrate, 2.8 % of the radio-labelled ricinoleoyl chains were found in triacylglycerol with one ricinoleoyl group, 12.4 % of the radioactivity was found in triacylglycerol-species with two ricinoleoyl groups and only trace amounts were associated with triacylglycerol consisting of three ricinoleoyl groups. If incubations with sn-2-[^{14}C] ricinoleoyl-diacylglycerol were performed in a 1:4 (mol:mol) mixture with non-radioactive diacylglycerol species containing one vernoloyl group, the distribution of radioactivity between different molecular species of triacylglycerol changed only marginally compared to incubations with just radioactive substrate. Only 1.3% of the added ^{14}C-labelled ricinoleoyl groups were metabolised into triacylglycerol species with one ricinoleoyl and one vernoloyl group. Similarly, only marginal changes in the radioactive triacylglycerol molecular species was seen in incubations where *sn*-2-[^{14}C]-ricinoleoyl-diacylglycerol was mixed with non-labelled divernoloyl-diacylglycerol. However, by adding unlabelled diricinoleoyl-phosphatidylcholine together with sn-2-[^{14}C]-ricinoleoyl-diacylglycerol the radioactivity metabolised into the different triacylglycerol species were substantially altered.

Only trace amounts of radioactivity were detected in triacylglycerol species with one ricinoleoyl chain whereas the radioactivity in triacylglycerol with two ricinoleoyl groups were doubled as compared to incubations with only sn-2-[^{14}C]-ricinoleoyl-diacylglycerol added.

Table 1. ***In vitro*** **synthesis of triacylglycerols in microsomal preparations of developing castor bean.** Aliqouts of microsomes (20 nmol PC) were lyophilised and substrate lipids were added in benzene solution: (A) 0.4 nmol [^{14}C]-DAG (7760 dpm/nmol) and where indicated 1.6 nmol unlabelled DAG; (B) 0.4 nmol [^{14}C]-DAG (7760 dpm/nmol) and 5 nmol unlabelled di-ricinoleoyl-PC and (C) 0.25nmol [^{14}C]-PC (4000 dpm/nmol) and 5 nmol unlabelled DAG. The benzene was evaporated by N_2 and 0.1 ml of 50 mM potassium phosphate was added, thoroughly mixed and incubated at 30 °C for (A) 20 min; (B) and (C) 30 min. Assays were terminated by extraction of the lipids in chloroform (Dahlqvist et al. 1998). The lipids were then separated by thin layer chromatography on silica gel 60 plates (Merck, Darmstadt, Germany) in hexan/diethylether/acetic acid 35:70:1.5. The radioactive lipids were visualised and the radioactivity quantified on the plate by electronic autoradiography (Instant Imager, Packard, US). Results are presented as mean values of two experiments.

	Substrate added		mol% of added [^{14}C]-acyl group in TAG (1)				
	[^{14}C]- lipid (2)	unlabelled lipid (2)	1-OH-TAG	2-OH-TAG	1-OH-1-ver-TAG	1-OH-2-ver-TAG	3-OH-TAG
A	mono-[^{14}C]-ricinoleoyl-DAG	mono-ricinoleoyl-DAG	2.8	12.4	-	-	-
	mono-[^{14}C]-ricinoleoyl-DAG	mono-vernoleoyl –DAG	3.2	12.1	1.3	-	-
	mono-[^{14}C]-ricinoleoyl-DAG	di-vernoleoyl –DAG	4	10	0.5	1.2	-
B	mono-[^{14}C]-ricinoleoyl-DAG	di-ricinoleoyl -PC	0.3	24.8	-	-	-
C	mono-[^{14}C]-ricinoleoyl-PC	none	6.8	8.0	-	-	4.7
	mono-[^{14}C]-ricinoleoyl-PC	di-oleoyl-DAG	8.6	9.8	-	-	5.0
	mono-[^{14}C]-ricinoleoyl-PC	mono-ricinoleoyl-DAG	5.7	16.7	-	-	1.9
	mono-[^{14}C]-ricinoleoyl-PC	di-ricinoleoyl-DAG	4.5	9.4	-	-	9.5
	mono-[^{14}C]-ricinoleoyl-PC	mono-vernoleoyl-DAG	6.0	11.5	10.9	0.5	7.4
	mono-[^{14}C]-ricinoleoyl-PC	di-vernoleoyl-DAG	6.7	10.8	1.1	8.4	6.8

Radioactivity in different triacylglycerols (TAG) species formed. Abbreviations used: 1-OH-, mono-ricinoleoyl-; 2-OH, di-ricinoleoyl-; 3-OH-, triricinolcoyl; 1-OH-1-ver-, mono-ricinoleoly-monovernoleoyl-; 1-OH-2-ver-, mono-ricinoleoyl-divernoleoyl-. Radiolabclled DAG and PC were prepared enzymatically. The radiolabelled ricinoleoyl group is attached at the *sn-2*-position of the lipid and unlabelled oleoyl group at the *sn-1*-position. Unlabelled DAG with vernoleoyl- or ricinoleoyl chains were prepared by the action of TAG lipase (REF) on oil of *Euphorbia lagascae* or Castor bean, respectively. Synthetic di-ricinoleoyl-PC was kindly provided from Metapontum Agribios (Italy)

EXAMPLE 2

Transformation and expression of YNROORw gene in yeast

The yeast mutant (strain B 10280) disrupted in the YNR008w gene, was transformed with the single copy plasmid pFL39 having the PDAT-gene (YNR008w) under the control of the endogenous promotor region (583 bp 5' untranslated) inserted into the cloning cassette. The transformed yeast was pre-cultivated at 28 °C for 20 h in defined YNB medium without tryptophane added. Cells were harvested and re-suspended in minimal medium (Meesters eI al., 1996), supplemented with 16 g/l glycerol to the original volume of the growth culture. The culture was further incubated for 24 h after which cells were harvested by centrifugation. Microsomal fraction of the yeast was prepared as deseribed in Example 1 above and was incubated in the presence of sn-2-[^{14}C]-ricinoleoyl-phosphatidylcholine (Fig 1, lane 7). This experiment clearly shows that the PDAT activity could be restored by the expression of the YNR008w gene in the mutant yeast strain B10280 normally lacking the PDAT-activity.

The effect of the over-expression of the PDAT gene on the lipid accumulation was studied by transforming the wild-type yeast (strain SCY62) with a plasmid pJN92 containing the PDAT gene (YNR008w) under the control of a GAL1-promotor. The transformed yeast was then cultivated at 28 °C in defined YBN medium lacking uracil. The expression of the PDAT gene was induced by the addition of 2 % *(vlv)* galactose after 10 hours growth and was further incubated for 18 hours. The yeast cells were harvested and the lipid content of the yeast was analysed by thin layer chromatography and gas liquid chromatography. The results are presented in Table 2. The total lipid content in the yeast with the over-expressed PDAT was 1.3 fold higher that in the control yeast transformed with an empty plasmid pJN92. The expression of the PDAT gene had no effect on the growth rate as determined by optical density measurements. The elevated lipid content in the yeast transformed with PDAT as compared to the control yeast can be totally accounted for by an 80 % increase seen in the triacylglycerol content. The levels of the polar lipids and sterol esters were not significantly effected by the over-expression of the PDAT gene. Hence, these results clearly demonstrate the use of the PDAT gene in increasing the oil content in transgenic organisms.

Table 2. ***Lipid content in yeast with overexpressed PDAT gene.***
Yeast transformed with the PDAT gene under the control of a gal promotor in a xxxx plasmid were grown in YNB medium at 28 °C. After after 10 hours growth the expression of the PDAT gene was induced by the addition of 2 % (v/v) galactose. Cells were harvested after additional 18 hours growth and analysed for its lipid content. Yeast transformed with an empty plasmid grown under identical conditions was used as control. Lipids were extracted in chloroform, fractionated on TLC and quantified by GC analyses. The lipid content was measured as nmol fatty acids (FA) per mg yeast (dry weight).

	PDAT-transformed yeast		Control yeast	
	nmol FA/mg	(%)	nmol FA/mg	(%)
Polar lipids	60	32	60	42
Sterol esters	17	9	20	14
Triacylglycerol	105	56	58	40
Other lipids	6	3	6	4
Sum	188		144	

Claims

1. An enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol forming triacylglycerol.

2. A recombinant DNA molecule encoding the enzyme of claim 1.

3. A recombinant DNA molecule according to claim 2 encoding an enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol, wherein said enzyme comprises an amino acid sequence set forth in SEQ ID NO. 2, or fragments thereof, or is encoded by the genes set forth in SEQ ID NO. 1, 3, 4, 9 or 10, or is encoded by the fullenght genes corresponding to the partial nucleotide sequences set forth if SEQ ID NO. 5, 6, 7, 8, and 11 or sequences 40% or more identical to said sequences.

4. The DNA molecule of claim 2 or 3, wherein said enzyme is designated as phospholipid:diacylglycerol acyltransferease.

5. A vector comprising a DNA molecule of claim 2 or 3.

6. A vector of claim 5 further comprising a selectable marker gene.

7. A host cell containing a DNA molecule of claim 2 or 3.

8. The host cell of claim 7 which is a plant cell or yeast cell.

9. A process for the production of transgenic yeast cells, plant cells or plants comprising a) transforming a DNA moleclule of claims 2 or 3 into plant cells or plants; and b) selecting of transformed plant cells or plants having an altered biosynthetic pathway in the production of triacylglycerol.

10. A process of claim 9 wherein the altered biosynthetic pathway is characterised by an increased or altered oil content.

11. A process of claim 9 wherein the altered biosynthetic pathway is characterised by the prevention of accumulation of undesirable fatty acids in the membrane lipids.

12. A protein encoded by a DNA molecule according to claim 2 or 3 or a functional fragment thereof.

13. A method for production of triacylglycerol, comprising enzymatic catalysation by an enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol forming triacylglycerol.

14. A method according to claim 13, comprising enzymatic catalysation by a protein or functional fragment according to claim 12.

Abstract

The present invention relates to a novel class of enzyme catalysing the transfer of fatty acids from phospholipids to diacylglycerol in the biosynthetic pathway for the production of triacylglycerol. Further, the invention relates to the isolation, identification, characterization and use of recombinant DNA molecules encoding this class of enzymes.

SEQUENCE LISTING

<110> Stymne Dr., Sten
<120> RECOMBINANT DNA MOLECULES ENCODING A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE PRODUCTION OF TRIACYLGLYCEROL
<130> stymne
<140>
<141>
<160> 11
<170>
<210> 1
<211> 1986
<212> genomic DNA
<213> Saccharomyces cerevisiae
<220>
<221> CDS
<222> (1)..(1983)
<400> 1

```
atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct    48
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
  1               5                  10                  15

gat gaa aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga    96
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
             20                  25                  30

aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt    144
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
         35                  40                  45

att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg    192
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg
     50                  55                  60

aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg    240
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
 65                  70                  75                  80

att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt    288
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe
                 85                  90                  95

ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt    336
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe
            100                 105                 110
```

gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt 384
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
115 120 125

ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac 432
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn
130 135 140

tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt 480
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly
145 150 155 160

aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta 528
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val
165 170 175

atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att 576
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile
180 185 190

gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg 624
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp
195 200 205

gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg 672
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
210 215 220

ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac 720
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn
225 230 235 240

ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc 768
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
245 250 255

gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att 816
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile
260 265 270

ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt 864
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
275 280 285

gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag 912
Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
290 295 300

gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta 960
Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
305 310 315 320

att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg 1008
Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
325 330 335

gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt 1056
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
340 345 350

aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc 1104
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
355 360 365

gct cca aag gca gtt cca gct cta att agt ggt gaa atg aaa gat acc 1152
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
370 375 380

att caa tta aat acg tta gcc atg tat ggt ttg gaa aag ttc ttc tca 1200
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
385 390 395 400

aga att gag aga gta aaa atg tta caa acg tgg ggt ggt ata cca tca 1248
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
405 410 415

atg cta cca aag gga gaa gag gtc att tgg ggg gat atg aag tca tct 1296
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser
420 425 430

tca gag gat gca ttg aat aac aac act gac aca tac ggc aat ttc att 1344
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
435 440 445

cga ttt gaa agg aat acg agc gat gct ttc aac aaa aat ttg aca atg 1392
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
450 455 460

aaa gac gcc att aac atg aca tta tcg ata tca cct gaa tgg ctc caa 1440
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
465 470 475 480

```
aga aga gta cat gag cag tac tcg ttc ggc tat tcc aag aat gaa gaa   1488
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
                485                 490                 495

gag tta aga aaa aat gag cta cac cac aag cac tgg tcg aat cca atg   1536
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
            500                 505                 510

gaa gta cca ctt cca gaa gct ccc cac atg aaa atc tat tgt ata tac   1584
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
        515                 520                 525

ggg gtg aac aac cca act gaa agg gca tat gta tat aag gaa gag gat   1632
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
    530                 535                 540

gac tcc tct gct ctg aat ttg acc atc gac tac gaa agc aag caa cct   1680
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
545                 550                 555                 560

gta ttc ctc acc gag ggg gac gga acc gtt ccg ctc gtg gcg cat tca   1728
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
                565                 570                 575

atg tgt cac aaa tgg gcc cag ggt gct tca ccg tac aac cct gcc gga   1776
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
            580                 585                 590

att aac gtt act att gtg gaa atg aaa cac cag cca gat cga ttt gat   1824
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
        595                 600                 605

ata cgt ggt gga gca aaa agc gcc gaa cac gta gac atc ctc ggc agc   1872
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
    610                 615                 620

gcg gag ttg aac gat tac atc ttg aaa att gca agc ggt aat ggc gat   1920
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
625                 630                 635                 640

ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag   1968
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
                645                 650                 655
```

```
atg ccc ttc cca atg taa                                 1986
Met Pro Phe Pro Met
            660
```

<210> 2
<211> 661
<212> PRT
<213> Saccharomyces cerevisiae

<400> 2

```
Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
1               5                   10                  15
Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
            20                  25                  30
Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
        35                  40                  45
Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg
    50                  55                  60
Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
65                  70                  75                  80
Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe
                85                  90                  95
Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe
            100                 105                 110
Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
        115                 120                 125
Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn
    130                 135                 140
Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly
145                 150                 155                 160
Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val
                165                 170                 175
Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile
            180                 185                 190
Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp
        195                 200                 205
Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
    210                 215                 220
Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn
225                 230                 235                 240
Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
                245                 250                 255
Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile
            260                 265                 270
Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
        275                 280                 285
Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
    290                 295                 300
Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
305                 310                 315                 320
```

Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
                325   330   335
Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
            340   345   350
Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
        355   360   365
Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
    370   375   380
Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
385   390   395   400
Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
                405   410   415
Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser
            420   425   430
Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
        435   440   445
Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
    450   455   460
Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
465   470   475   480
Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
                485   490   495
Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
            500   505   510
Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
        515   520   525
Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
    530   535   540
Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
545   550   555   560
Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
                565   570   575
Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
            580   585   590
Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
        595   600   605
Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
    610   615   620
Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
625   630   635   640
Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
                645   650   655
Met Pro Phe Pro Met
            660

<210> 3
<211> 2312
<212> genomic DNA
<213> Schizosaccharomyces pombe
<400> 3

ATGGCGTCTT CCAAGAAGAG CAAAACTCAT AAGAAAAAGA AAGAAGTCAA 50

ATCTCCTATC GACTTACCAA ATTCAAAGAA ACCAACTCGC GCTTTGAGTG 100

AGCAACCTTC AGCGTCCGAA ACACAATCTG TTTCAAATAA ATCAAGAAAA 150

TCTAAATTTG GAAAAAGATT GAATTTTATA TTGGGCGCTA TTTTGGGAAT 200

ATGCGGTGCT TTTTTTTTCG CTGTTGGAGA CGACAATGCT GTTTTCGACC 250

CTGCTACGTT AGATAAATTT GGGAATATGC TAGGCTCTTC AGACTTGTTT 300

GATGACATTA AAGGATATTT ATCTTATAAT GTGTTTAAGG ATGCACCTTT 350

TACTACGGAC AAGCCTTCGC AGTCTCCTAG CGGAAATGAA GTTCAAGTTG 400

GTCTTGATAT GTACAATGAG GGATATCGAA GTGACCATCC TGTTATTATG 450

GTTCCTGGTG TTATCAGCTC AGGATTAGAA AGTTGGTCGT TTAATAATTG 500

CTCGATTCCT TACTTTAGGA AACGTCTTTG GGGTAGCTGG TCTATGCTGA 550

AGGCAATGTT CCTTGACAAG CAATGCTGGC TTGAACATTT AATGCTTGAT 600

AAAAAAACCG GCTTGGATCC GAAGGGAATT AAGCTGCGAG CAGCTCAGGG 650

GTTTGAAGCA GCTGATTTTT TTATCACGGG CTATTGGATT TGGAGTAAAG 700

TAATTGAAAA CCTTGCTGCA ATTGGTTATG AGCCTAATAA CATGTTAAGT 750

GCTTCTTACG ATTGGCGGTT ATCATATGCA AATTTAGAGG AACGTGATAA 800

ATATTTTTCA AAGTTAAAAA TGTTCATTGA GTACAGCAAC ATTGTACATA 850

AGAAAAAGGT AGTGTTGATT TCTCACTCCA TGGGTTCACA GGTTACGTAC 900

TATTTTTTTA AGTGGGTTGA AGCTGAGGGC TACGGAAATG GTGGACCGAC 950

TTGGGTTAAT GATCATATTG AAGCATTTAT AAATGTGAGT CTCGATGGTT 1000

GTTTGACTAC GTTTCTAACT TTTGAATAGA TATCGGGATC TTTGATTGGA 1050

GCACCCAAAA CAGTGGCAGC GCTTTTATCG GGTGAAATGA AAGATACAGG 1100

TATTGTAATT ACATTAAACA TGTTAATATT TAATTTTTGC TAACCGTTTT 1150

AAGCTCAATT GAATCAGTTT TCGGTCTATG GGTAAGCAAT AAATTGTTGA 1200

GATTTGTTAC TAATTTACTG TTTAGTTTGG AAAAATTTTT TTCCCGTTCT 1250

GAGGTATATT CAAAAATACA AATGTGCTCT ACTTTTTCTA ACTTTTAATA 1300

GAGAGCCATG ATGGTTCGCA CTATGGGAGG AGTTAGTTCT ATGCTTCCTA 1350

AAGGAGGCGA TGTTGTATGG GGAAATGCCA GTTGGGTAAG AAATATGTGC 1400

TGTTAATTTT TTATTAATAT TTAGGCTCCA GATGATCTTA ATCAAACAAA 1450
TTTTTCCAAT GGTGCAATTA TTCGATATAG AGAAGACATT GATAAGGACC 1500
ACGATGAATT TGACATAGAT GATGCATTAC AATTTTTAAA AAATGTTACA 1550
GATGACGATT TTAAAGTCAT GCTAGCGAAA AATTATTCCC ACGGTCTTGC 1600
TTGGACTGAA AAAGAAGTGT TAAAAAATAA CGAAATGCCG TCTAAATGGA 1650
TAAATCCGCT AGAAGTAAGA ACATTAAAGT TACTAAATTA TACTAACCCA 1700
AATAGACTAG TCTTCCTTAT GCTCCTGATA TGAAAATTTA TTGCGTTCAC 1750
GGGGTCGGAA AACCAACTGA GAGAGGTTAT TATTATACTA ATAATCCTGA 1600
GGGGCAACCT GTCATTGATT CCTCGGTTAA TGATGGAACA AAAGTTGAAA 1850
ATGTGAGAGA ATTTATGTTT CAAACATTCT ATTAACTGTT TTATTAGGGT 1900
ATTGTTATGG ATGATGGTGA TGGAACTTTA CCAATATTAG CCCTTGGTTT 1950
GGTGTGCAAT AAAGTTTGGC AAACAAAAAG GTTTAATCCT GCTAATACAA 2000
GTATCACAAA TTATGAAATC AAGCATGAAC CTGCTGCGTT TGATCTGAGA 2050
GGAGGACCTC GCTCGGCAGA ACACGTCGAT ATACTTGGAC ATTCAGAGCT 2100
AAATGTATGT TCATTTTACC TTACAAATTT CTATTACTAA CTCTTGAAAT 2150
AAGGAAATTA TTTTAAAAGT TTCATCAGGC CATGGTGACT CGGTACCAAA 2200
CCGTTATATA TCAGATATCC AGTACGGACA TAAGTTTTGT AGATTGCAAT 2250
TAACTAACTA ACCGAACAGG GAAATAATAA ATGAGATAAA TCTCGATAAA 2300
CCTAGAAATT AA 2312

<210> 4
<211> 3685
<212> genomic DNA
<213> Arabidopsis thaliana
<400> 4

```
ATGCCCCTTA TTCATCGGAA AAAGCCGACG GAGAAACCAT CGACGCCGCC     50
ATCTGAAGAG GTGGTGCACG ATGAGGATTC GCAAAAGAAA CCACACGAAT    100
CTTCCAAATC CCACCATAAG AAATCGAACG GAGGAGGGAA GTGGTCGTGC    150
ATCGATTCTT GTTGTTGGTT CATTGGGTGT GTGTGTGTAA CCTGGTGGTT    200
TCTTCTCTTC CTTTACAACG CAATGCCTGC GAGCTTCCCT CAGTATGTAA    250
CGGAGCGAAT CACGGGTCCT TTGCCTGACC CGCCCGGTGT TAAGCTCAAA    300
AAAGAAGGTC TTAAGGCGAA ACATCCTGTT GTCTTCATTC CTGGGATTGT    350
CACCGGTGGG CTCGAGCTTT GGGAAGGCAA ACAATGCGCT GATGGTTTAT    400
TTAGAAAACG TTTGTGGGGT GGAACTTTTG GTGAAGTCTA CAAAAGGTGA    450
GCTCAACAAT TCTCACTCTT CCTTTATATT GGGATTTGGA TTGGATCTGA    500
TGAGATCACG CACTTGTTGC TTCTTCAACA TCACTCAAAC TTTAATTCCA    550
TGTTTGTCTG TCTTACTCTT TACTTTTTTT TTTTTTTGAT GTGAAACGCT    600
ATTTTCTTAA GAGACTATTT CTGTATGTGT AAGGTAAGCG TTCCAAGGAC    650
GTAATTGGCT TGGACTATTT CTGTTTGATT GTTAACTTTA GGATATAAAA    700
TAGCTGCCTT GGAATTTCAA GTCATCTTAT TGCCAAATCT GTTGCTAGAC    750
ATGCCCTAGA GTCCGTTCAT AACAAGTTAC TTCCTTTACT GTCGTTGCGT    800
GTAGATTTAG CTTTGTGTAG CGTATAATGA AGTAGTGTTT TATGTTTTGT    850
TGGGAATAGA GAAGTTCTAA CTACATCTGT GGAAAGTGTG TTCAGGCTGT    900
GATAGAGGAC TGTTGCTTTA TTATTCAACT ATGTATATGT GTAATTAAAG    950
CTAGTTCCTT TTTGATCTTT CAGCTCAATG TGCTTTTCTC AATTTTTTTC   1000
TCAATTTCAA AGTTTCACAT CGAGTTTATT CACATGTCTT GAATTTCGTC   1050
CATCCTCGTT CTGTTATCCA GCTTTGAACT CCTCCCGACC CTGCTATGGA   1100
TATATTAAAA AAAAAGTGTT TTGTGGGTTG CATCTTTGTT ACGATCTGCA   1150
TCTTCTTCTT TCGGCTCAGT GTTCATGTTT TTGCTATGGT AGAGATGGGC   1200
AATGTTATTG TTGATGGTAA CAGTGGTATA GTTGATAGTA TCTTAACTAA   1250
TCAATTATCT CTTTGATTCA GGCCTCTATG TTGGGTGGAA CACATGTCAC   1300
TTGACAATGA AACTGGGTTG GATCCAGCTG GTATTAGAGT TCGAGCTGTA   1350
```

TCAGGACTCG TGGCTGCTGA CTACTTTGCT CCTGGCTACT TTGTCTGGGC 1400

AGTGCTGATT GCTAACCTTC CACATATTGG ATATGAAGAG AAAAATATGT 1450

ACATGGCTGC ATATGACTGG CGGCTTTCGT TTCAGAACAC AGAGGTTCTT 1500

TTCTCATCGT TCTTTCTATT ATTCTGTTCC ATGTTACGTT TCTTTCTTCA 1550

TTACTTAAGG CTTAAATATG TTTCATGTTG AATTAATAGG TACGTGATCA 1600

GACTCTTAGC CGTATGAAAA GTAATATAGA GTTGATGGTT TCTACCAACG 1650

GTGGAAAAAA AGCAGTTATA GTTCCGCATT CCATGGGGGT CTTGTATTTT 1700

CTACATTTTA TGAAGTGGGT TGAGGCACCA GCTCCTCTGG GTGGCGGGGG 1750

TGGGCCAGAT TGGTGTGCAA AGTATATTAA GGCGGTGATG AACATTGGTG 1800

GACCATTTCT TGGTGTTCCA AAAGCTGTTG CAGGGCTTTT CTCTGCTGAA 1850

GCAAAGGATG TTGCAGTTGC CAGGTATTGA ATATCTGCTT ATACTTTTGA 1900

TGATCAGAAC CTTGGCTCTG GAACTCAAAG TTATTCTACT AAATATCAAT 1950

TCTAATAACA TTGCTATATT ATCGCTGCAA CTGACATTGG TTGATTATTT 2000

TTGCTGCTTA TGTAACTGAA ACTCTCTTGA GATTAGACAA ATGATGAATT 2050

GATAATTCTT ACGCATTGCT CTGTGATGAC CAGTTTCTTA GCTTCGACGA 2100

TAACATTTGT CATACTGTCT TTTGGAGGGC ATTGAATTTT GCTATGGAAA 2150

GCGCTGGAGC TTCCATGCTT GCATTCTTTA CCAATTAGCG TTATTCTGCT 2200

TCTTTCAATT TTCTTGTATA TGCATCTATG GTCTTTTATT TCTTCTTAAT 2250

TAAAGACTCG TTGGATTAGT TGCTCTATTA GTCACTTGGT TCCTTAATAT 2300

AGAACTTTAC TTTCTTCGAA AATTGCAGAG CGATTGCCCC AGGATTCTTA 2350

GACACCGATA TATTTAGACT TCAGACCTTG CAGCATGTAA TGAGAATGAC 2400

ACGCACATGG GACTCAACAA TGTCTATGTT ACCGAAGGGA GGTGACACGA 2450

TATGGGGCGG GCTTGATTGG TCACCGGAGA AAGGCCACAC CTGTTGTGGG 2500

AAAAAGCAAA AGAACAACGA AACTTGTGGT GAAGCAGGTG AAAACGGAGT 2550

TTCCAAGAAA AGTCCTGTTA ACTATGGAAG GATGATATCT TTTGGGAAAG 2600

```
CTATAGATCT ACTACATTAT GTTGCTCCTA AGATGATGGC GCGTGGTGCC     2900
GCTCATTTCT CTTATGGAAT TGCTGATGAT TTGGATGACA CCAAGTATCA     2950
AGATCCCAAA TACTGGTCAA ATCCGTTAGA GACAAAGTAA GTGATTTCTT     3000
GATTCCAACT GTATCCTTCG TCCTGATGCA TTATCAGTCT TTTTGTTTTC     3050
GGTCTTGTTG GATATGGTTT TCAGCTCAAA GCTTACAAAG CTGTTTCTGA     3100
GCCTTTCTCA AAAAGGCTTG CTCAGTAATA TTGAGGTGCT AAAGTTGATA     3150
CATGTGACTC TTGCTTATAA ATCCTCCGTT TGGTTTGTTC TGCTTTTTCA     3200
GATTACCGAA TGCTCCTGAG ATGGAAATCT ACTCATTATA CGGAGTGGGG     3250
ATACCAACGG AACGAGCATA CGTATACAAG CTTAACCAGT CTCCCGACAG     3300
TTGCATCCCC TTTCAGATAT TCACTTCTGC TCACGAGGAG GACGAAGATA     3350
GCTGTCTGAA AGCAGGAGTT TACAATGTGG ATGGGGATGA AACAGTACCC     3400
GTCCTAAGTG CCGGGTACAT GTGTGCAAAA GCGTGGCGTG GCAAGACAAG     3450
ATTCAACCCT TCCGGAATCA AGACTTATAT AAGAGAATAC AATCACTCTC     3500
CGCCGGCTAA CCTGTTGGAA GGGCGCGGGA CGCAGAGTGG TGCCCATGTT     3550
GATATCATGG GAAACTTTGC TTTGATCGAA GATATCATGA GGGTTGCCGC     3600
CGGAGGTAAC GGGTCTGATA TAGGACATGA CCAGGTCCAC TCTGGCATAT     3650
TTGAATGGTC GGAGCGTATT GACCTGAAGC TGTGA                     3685
```

<210> 5
<211> 402
<212> cDNA
<213> Arabidopsis thaliana
<220>
<221> CDS
<222> (120)..(402)
<400> 5

```
AGAAACAGCTCTTTGTCTCT CTCGACTGATCTAACAATCC CTAATCTGTGTTCTAAATTC   60

CTGGACGAGATTTGACAAAG TCCGTATAGCTTAACCTGGT TTAATTTCAAGTGACAGAT    119

ATG CCC CTT ATT CAT CGG AAA AAG CCG ACG GAG AAA CCA TCG ACG CCG  167
Met Pro Leu Ile His Arg Lys Lys Pro Thr Glu Lys Pro Ser Thr Pro

CCA TCT GAA GAG GTG GTG CAC GAT GAG GAT TCG CAA AAG AAA CCA CAC  215
Pro Ser Glu Glu Val Val His Asp Glu Asp Ser Gln Lys Lys Pro His

GAA TCT TCC AAA TCC CAC CAT AAG NAA TCG AAC GGA GGA GGG AAG TGG  263
Glu Ser Ser Lys Ser His His Lys ??? Ser Asn Gly Gly Gly Lys Trp

TCG TGC ATC GAT TCT TGT TGT TGG TTC ATT GGG TGT GTG TGT GTA ACC  311
Ser Cys Ile Asp Ser Cys Cys Trp Phe Ile Gly Cys Val Cys Val Thr

TGG TGG TTT CTT CTC TTC CTT TAC AAC GCA ATG CCT GCG AGC TTC CCT  359
Trp Trp Phe Leu Leu Phe Leu Tyr Asn Ala Met Pro Ala Ser Phe Pro

CAG TAT GTA ACG GAG CCG AAT CAC GNG TCC TTT GCC TTA CCC G        402
Gln Tyr Val Thr Glu Pro Asn His ??? Ser Phe Ala Leu Pro
```

```
<210> 6
<211> 643
<212> cDNA
<213> Zea mays
<220>
<221> CDS
<222> (1)..(402)
<400> 6

CGG GAG AAA ATA GCT GCT TTG AAG GGG GGT GTT TAC TTA GCC GAT GGT   48
Arg Glu Lys Ile Ala Ala Leu Lys Gly Gly Val Tyr Leu Ala Asp Gly
 1               5                  10                  15

GAT GAA ACT GTT CCA GTT CTT AGT GCG GGC TAC ATG TGT GCG AAA GGA   96
Asp Glu Thr Val Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Gly
            20                  25                  30

TGG CGT GGC AAA ACT CGT TTC AGC CCT GCC GGC AGC AAG ACT TAC GTG   144
Trp Arg Gly Lys Thr Arg Phe Ser Pro Ala Gly Ser Lys Thr Tyr Val
        35                  40                  45

AGA GAA TAC AGC CAT TCG CCA CCC TCT ACT CTC CTG GAA GGC AGG GGC   192
Arg Glu Tyr Ser His Ser Pro Pro Ser Thr Leu Leu Glu Gly Arg Gly
    50                  55                  60

ACC CAG AGC GGT GCA CAT GTT GAT ATA ATG GGG AAC TTT GCT CTA ATT   240
Thr Gln Ser Gly Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile
65                  70                  75                  80

GAG GAC GTC ATC AGA ATA GCT GCT GGG GCA ACC GGT GAG GAA ATT GGT   288
Glu Asp Val Ile Arg Ile Ala Ala Gly Ala Thr Gly Glu Glu Ile Gly
                85                  90                  95

GGC GAT CAG GTT TAT TCA GAT ATA TTC AAG TGG TCA GAG AAA ATC AAA   336
Gly Asp Gln Val Tyr Ser Asp Ile Phe Lys Trp Ser Glu Lys Ile Lys
            100                 105                 110

TTG AAA TTG TAA  CCTATGGGAA GTTAAAGAAG TGCCGACCCG TTTATTGCGTTCC   391

Leu Lys Leu
        115

AAAGTGTCCT GCCTGAGTGC AACTCTGGAT TTTGCTTAAA TATTGTAATT TTTCACGC   449

TTCATTCGTC CCTTTGTCAA ATTTACATTT GACAGGACGC CAATGCGATA CGATGTTG   507

TACCGCTATT TTCAGCATTG TATATTAAAC TGTACAGGTG TAAGTTGCAT TTGCCAGC   565

TGAAATTGTG TAGTCGTTTT CTTTACGATT TAATANCAAG TGGCGGAGCA GTGCCCCA   623

AGCNAAAAAA AAAAAAAAAA                                             643
```

<210> 7
<211> 115
<212> PRT
<213> Zea mays
<400> 7

Arg	Glu	Lys	Ile	Ala	Ala	Leu	Lys	Gly	Gly	Val	Tyr	Leu	Ala	Asp	Gly
1				5					10					15	
Asp	Glu	Thr	Val	Pro	Val	Leu	Ser	Ala	Gly	Tyr	Met	Cys	Ala	Lys	Gly
			20					25					30		
Trp	Arg	Gly	Lys	Thr	Arg	Phe	Ser	Pro	Ala	Gly	Ser	Lys	Thr	Tyr	Val
		35					40					45			
Arg	Glu	Tyr	Ser	His	Ser	Pro	Pro	Ser	Thr	Leu	Leu	Glu	Gly	Arg	Gly
	50					55					60				
Thr	Gln	Ser	Gly	Ala	His	Val	Asp	Ile	Met	Gly	Asn	Phe	Ala	Leu	Ile
65					70					75					80
Glu	Asp	Val	Ile	Arg	Ile	Ala	Ala	Gly	Ala	Thr	Gly	Glu	Glu	Ile	Gly
				85					90					95	
Gly	Asp	Gln	Val	Tyr	Ser	Asp	Ile	Phe	Lys	Trp	Ser	Glu	Lys	Ile	Lys
			100					105					110		
Leu	Lys	Leu													
		115													

<210> 8
<211> 616
<212> cDNA
<213> Neurospora crassa
<400> 8

```
GGTGGCGAAG ACGANGGCGG AAGTTGGAGG CTAACGAGAA TGACNCTCGG        50
AGATGGATCT ACCCTCTAGA GACACGACTA CCNTTGCACC CAGCCTCAAG       100
GTNTACNGTT TNTATGGGTA GGAAGCCGAC GGAGCGAGCC TACATCTATC       150
TGCCGCCCGA TCCCGGGACG ACAACGCATC TTTAGATGAC GATCGATACG       200
ACTTTGACTN AGGGGCACAT TGACCACGGT GTGATTTTGG GCGAAGGCGA       250
TGGCACAGTG AACCTTATGA GTTTGGGGTA CCTGTGCAAT AAGGGGTGGA       300
AAATGAAGAG ATACAATCCT GCGGGCTCAA AAATAACCGT GGTCGAGATG       350
CCGCATGAAC CAGAACGGTT CAATCCGAGA GGAGGGCCGA ATACGGCGGA       500
CTTAAATATG TAGAAAAGGT TGAAATTTAT GAAGAGTAAT TAAATACGGC       550
ACATAGGTTA CTCAATAGTA TGACTAATTA AAAAAAAATT TTTTTCTAA        600
AAAAAAAAAA AAAAAA                                            616
```

<210> 9
<211> 1562
<212> genomic DNA
<213> Arabidopsis thaliana
<400> 9

ATGAAAAAAA TATCTTCACA TTATTCGGTA GTCATAGCGA TACTCGTTGT 50

GGTGACGATG ACCTCGATGT GTCAAGCTGT GGGTAGCAAC GTGTACCCTT 100

TGATTCTGGT TCCAGGAAAC GGAGGTAACC AGCTAGAGGT ACGGCTGGAC 150

AGAGAATACA AGCCAAGTAG TGTCTGGTGT AGCAGCTGGT TATATCCGAT 200

TCATAAGAAG AGTGGTGGAT GGTTTAGGCT ATGGTTCGAT GCAGCAGTGT 250

TATTGTCTCC CTTCACCAGG TGCTTCAGCG ATCGAATGAT GTTGTACTAT 300

GACCCTGATT TGGATGATTA CCAAAATGCT CCTGGTGTCC AAACCCGGGT 350

TCCTCATTTC GGTTCGACCA AATCACTTCT ATACCTCGAC CCTCGTCTCC 400

GGTTAGTACT TTCCAAGATA TATCATTTTG GGACATTTGC ATAATGAACA 450

AAATAGACAT AAATTTGGGG GATTATTGTT ATATCAATAT CCATTTATAT 500

GCTAGTCGGT AATGTGAGTG TTATGTTAGT ATAGTTAATG TGAGTGTTAT 550

GTGATTTTCC ATTTTAAATG AAGCTAGAAA GTTGTCGTTT AATAATGTTG 600

CTATGTCATG AGAATTATAA GGACACTATG TAAATGTAGC TTAATAATAA 650

GGTTTGATTT GCAGAGATGC CACATCTTAC ATGGAACATT TGGTGAAAGC 700

TCTAGAGAAA AAATGCGGGT ATGTTAACGA CCAAACCATC CTAGGAGCTC 750

CATATGATTT CAGGTACGGC CTGGCTGCTT CGGGCCACCC GTCCCGTGTA 800

GCCTCACAGT TCCTACAAGA CCTCAAACAA TTGGTGGAAA AAACTAGCAG 850

CGAGAACGAA GGAAAGCCAG TGATACTCCT CTCCCATAGC CTAGGAGGAC 900

TTTTCGTCCT CCATTTCCTC AACCGTACCA CCCCTTCATG GCGCCGCAAG 950

TACATCAAAC ACTTTGTTGC ACTCGCTGCG CCATGGGGTG GGACGATCTC 1000

TCAGATGAAG ACATTTGCTT CTGGCAACAC ACTCGGTGTC CCTTTAGTTA 1050

ACCCTTTGCT GGTCAGACGG CATCAGAGGA CCTCCGAGAG TAACCAATGG 1100

CTACTTCCAT CTACCAAAGT GTTTCACGAC AGAACTAAAC CGCTTGTCGT 1150

AACTCCCCAG GTTAACTACA CAGCTTACGA GATGGATCGG TTTTTTGCAG 1200

ACATTGGATT CTCACAAGGA GTTGTGCCTT ACAAGACAAG AGTGTTGCCT 1250

```
TTAACAGAGG AGCTGATGAC TCCGGGAGTG CCAGTCACTT GCATATATGG    1300
GAGAGGAGTT GATACACCGG AGGTTTTGAT GTATGGAAAA GGAGGATTCG    1350
ATAAGCAACC AGAGATTAAG TATGGAGATG GAGATGGGAC GGTTAATTTG    1400
GCGAGCTTAG CAGCTTTGAA AGTCGATAGC TTGAACACCG TAGAGATTGA    1450
TGGAGTTTCG CATACATCTA TACTTAAAGA CGAGATCGCA CTTAAAGAGA    1500
TTATGAAGCA GATTTCAATT ATTAATTATG AATTAGCCAA TGTTAATGCC    1550
GTCAATGAAT GA                                             1562
```

<210> 10
<211> 3896
<212> genomic DNA
<213> Arabidopsis thaliana
<400> 10

ATGGGAGCGA ATTCGAAATC AGTAACGGCT TCCTTCACCG TCATCGCCGT 50
TTTTTTCTTG ATTTGCGGTG GCCGAACTGC GGTGGAGGAT GAGACCGAGT 100
TTCACGCCGA CTACTCGAAG CTATCGGGTA TAATCATTCC GGGATTTGCG 150
TCGACGCAGC TACGAGCGTG GTCGATCCTT GACTGTCCAT ACACTCCGTT 200
GGACTTCAAT CCGCTCGACC TCGTATGGCT AGACACCACT AAGGTCCGTG 250
ATCTTCATTT CCTTCGCTCC TTATTCTGTC GGTCGAGTCA CTTGTTGATG 300
AATTCCAAGC GAAATATAGC AATGAAGCAT GTCTCGTCTC TCTTATTGAT 350
TCGTTCATTA GTCAACAGTG ACGCTTCTGA ATCTGAGTTT AGAGTCATAT 400
AAAACAGCTG ACTCGGCGAG TGTTTCCCAT CGCTTTTGGT TCGCTAAATG 450
TAGCGCAATG AATGTGTAAT TAGTCTGCGC TTTTTATTCA ACTAGATCTG 500
CAAGTTTTTC AGAGTGCTCA ATAGTAGTTA GAAAATGTTA GGTCATTTTA 550
CTTGTGCATT GTGATTCTTT TGGTTGTTGC TTACTGATCG ACGTGATGGA 600
TGGTTTACAG CTTCTTTCTG CTGTCAACTG CTGGTTTAAG TGTATGGTGC 650
TAGATCCTTA TAATCAAACA GACCATCCCG AGTGTAAGTC ACGGCCTGAC 700
AGTGGTCTTT CAGCCATCAC AGAATTGGAT CCAGGTTACA TAACAGGTAG 750
TTTCGGATTT TTCTTTCTTT TGAGTTTTCT TCAATTTGAT ATCATCTTGT 800
TGTGATATAA TATGGCTAAG TTCATTAATT TGGTCAATTT TCAGGTCCTC 850
TTTCTACTGT CTGGAAAGAG TGGCTTAAGT GGTGTGTTGA GTTTGGTATA 900
GAAGCAAATG CAATTGTCGC TGTTCCATAC GATTGGAGAT TGTCACCAAC 950
CAAATTGGAA GAGCGTGACC TTTACTTTCA CAAGCTCAAG TTAGTCCTTA 1000
TCAGGCTAAT GTCTTTTATC TTCTCTTTTT ATGTAAGATA AGCTAAGAGC 1050
TCTGGTCGTC TTCCTTTTTG CAGGTTGACC TTTGAAACTG CTTTAAAACT 1100
CCGTGGCGGC CCTTCTATAG TATTTGCCCA TTCAATGGGT AATAATGTCT 1150
TCAGATACTT TCTGGAATGG CTGAGGCTAG AAATTGCACC AAAACATTAT 1200
TTGAAGTGGC TTGATCAGCA TATCCATGCT TATTTCGCTG TTGGTACCGG 1250
CCTACTATCC TTAAGTTACC ATTTTATTTT TTCTCTAATT GGGGGAGTTA 1300
TGTTGTGACT TACTGGATTG AGCTCGATAC CTGATTTGTT GTTGATTTAG 1350
GAGCTCCTCT TCTTGGTTCT GTTGAGGCAA TCAAATCTAC TCTCTCTGGT 1400
GTAACGTTTG GCCTTCCTGT TTCTGAGGTG ACCTCTGACT TCTCTTTAGT 1450
TTTAAGTAGT TGATATCAAC CAGGTCTTAT AACTCACTGG ATTTTCCTTT 1500
TGAAAGTATT ACTTTTGTTA ATTGAACTGC TGTACGCGAT ATGGTATCTG 1550
TAGATCTTGA AGTGCTAGTT ATCAAAGAAC ATATTGTGGG TAGTATACCT 1600
GTCAGCGGCC TTAGCTAATA CAACCAAACC ACATGTACAC TGATTTAGTT 1650
TTCAGATTAT TATGGTAGAC TTTAAGTTGA GAAGAAACTT TGACTGAAAT 1700
CTTTTATTT TAATAGGCTA TGATTTGTTT ATTGAAATCA TGTGACATAT 1750
TGACATGCGC TTCTCATGTT TTTTGTTGGC AAGGCTTCAG GGAACTGCTC 1800
GGTTGTTGTC CAATTCTTTT GCGTCGTCAT TGTGGCTTAT GCCATTTTCA 1850
AAGAATTGCA AGGGTGATAA CACATTCTGG ACGCATTTTT CTGGGGGTGC 1900
TGCAAAGAAA GATAAGCGCG TATACCACTG TGATGAAGAG GAATATCAAT 1950
CAAAATATTC TGGCTGGCCG ACAAATATTA TTAACATTGA AATTCCTTCC 200
ACTAGCGGTT AGACTCTGTA TATGCAACTG TAACACTAAC AAAAGTTTCA 2050
CCAAGAATGT TCACTCTCAT ATTTCGTTCC TTTGATGTGT ATCCATCAGT 2100
TACAGAAACA GCTCTAGTCA ACATGACCAG CATGGAATGT GGCCTTCCCA 2150
CCCTTTTGTC TTTCACAGCC CGTGAACTAG CAGATGGGAC TCTTTTCAAA 2200
GCAATAGAAG ACTATGACCC AGATAGCAAG AGGATGTTAC ACCAGTTAAA 2250
GAAGTACGTA CCTTTCTTTG TGATAAGAAA TATTGCTCAT CGATCATCAC 2300
TTGCTGGCTT CTTGTACGTC AAATTGTTTT GTTTAAATCT CTATATCAAT 2350
TGTTCATATG CTTTGTCTTT CTTACTATAA GAAACAAGTA TAATCAGAAA 2400
CCTTATTATT GATTATCAGT TCTCTCCTTA TATTATGGAA TGTCTTTTTC 2450
GTTTACAGTT ATGAATGCAA AAGGGGGTAT TTTAGTTGAT TGATTCTCTC 2500
ATTCTCTAGT TTGTTTTGAC TAATAGCGTC AATTTTGTTT TTCTAGCAAA 2550
TCTTTGTGAA TTATATATAA CATGCTAACT ATACTTTTCA GGTTGTATCA 2600
TGATGACCCT GTTTTTAATC CTCTGACTCC TTGGGAGAGA CCACCTATAA 2650
AAAATGTATT TTGCATATAT GGTGCTCATC TAAAGACAGA GGTATGATGC 2700
ATTCTCAATA TCACATTATG CGTTGACTTT GTTATTATAT TCCCCATTTG 2750

GTTTGCAATA	TCTTTTTGAA	TTATGATTTA	TCTTCTCCCT	TGCATCTTAT	2800
GCTATTAAGC	GTTAAAGGTA	CTAAATGTAT	GAAGCTGTCT	GTCATAGGTT	2850
GGTTATTACT	TTGCCCCAAG	TGGCAAACCT	TATCCTGATA	ATTGGATCAT	2900
CACGGATATC	ATTTATGAAA	CTGAAGGTTC	CCTCGTGTCA	AGGTAATTTT	2950
CCGCAATGGC	AGAAGTAAAA	CAGGAAGGCA	AAGTCTTCTG	TATCAGTCTA	3000
GTGGCATGTT	ATCTCAGTTG	CATAAGCAAA	TTATTAAACA	ACTAAAATTT	3050
AAGTACTTTT	TTATCATTCC	TTTTGAGCTT	AGTGGATGAT	CAGTGGCTTA	3100
AACTGGGAAG	AGGTGTTGCA	TGAAACATGA	CACTTGTATC	AAAGATAACT	3150
AGCAAAACAA	AACTAACCCA	TTTCTGAATT	TCATATTATT	AGGAGTAGTC	3200
GTGCTTTTAA	AAAATTTGTT	TTAAGAAACC	GAAAAACTAG	TTCATATCTT	3250
GATTGTGCAA	TATCTGCAGG	TCTGGAACTG	TGGTTGATGG	GAACGCTGGA	3300
CCTATAACTG	GGGATGAGAC	GGTAAGCTCA	GAAGTTGGTT	TTGAAATTAT	3350
CTTCTTGCAA	ACTACTGAAG	ACTAAGATAA	TACTTGCTTC	TGGAACACTG	3400
CTTGCTATGT	TCTCTAGTAC	ACTGCAATAT	TGACTCTCCG	CTACTTTTAT	3450
TGATTATGAA	ATTGATCTCT	TATAGGTACC	CTATCATTCA	CTCTCTTGGT	3500
CCAAGAATTG	GCTCGGACCT	AAAGTTAACA	TAACAATGGC	TCCCCAGGTA	3550
CTCTTTTTTA	GTTCCTCACC	TTATATAGAT	CAAACTTTAA	GTGTACTTTT	3600
CTGGTTATGT	GTTGATTTAC	CTCCAATTTG	TTCTTTCTAA	AAATCATATA	3650
TCTCTGTACT	CCTCAAGAAC	TTGTATTAAT	CTAAACGAGA	TTCTCATTGG	3700
GAAAATAAAA	CAACAGCCAG	AACACGATGG	AAGCGACGTA	CATGTGGAAC	3750
TAAATGTTGA	TCATGAGCAT	GGGTCAGACA	TCATAGCTAA	CATGACAAAA	3800
GCACCAAGGG	TTAAGTACAT	AACCTTTTAT	GAAGACTCTG	AGAGCATTCC	3850
GGGGAAGAGA	ACCGCAGTCT	GGGAGCTTGA	TAAAAGTGGG	TATTAA	3896

<210> 11
<211> 709
<212> cDNA
<213> tomato
<400> 11

```
CTGGGGCCAA AAGTGAACAT AACAAGGACA CCACAGTCAG AGCATGATGT    50
TCAGATGTAC AAGTGCATCT AAATATAGAG CATCAACATG GTGAAGATAT   100
CATTCCCAAT ATGACAAAGT TACCTACAAT GAAGTACATA ACCTATTATG   150
AGGATTCTGA AAGTTTTCCA GGGACAAGAA CAGCAGTTTG GGAGCTTGAT   200
AAAGCAAATC ACAGGAACAT TGTCAGATCT CCAGCTTTGA TGCGGGAGCT   250
GTGGCTTGAG ATGTGGCATG ATATTCATCC TGATAAAAAG TCCAAGTTTG   300
TTACAAAAGG TGGTGTCTGA TCCTCACTAT TTTCTTCTAT AAATGTTTGA   350
GTTTGTATTG ACATTGTAAG TATTGCAACA AAAAGCAAAG CGTGGGCCTC   400
TGAGGGATGA GGACTGCTAT TGGGATTACG GGAAAGCTCG ATGTGCATGG   450
GCTGAACATT GTGAATACAG GTTAGAATAT TCAAATTATA TTTTGCAAAA   500
TATTCTCTTT TTGTGTATTT AGGCCACCTT TCCCCGGTCA CAACGATGCA   550
GATATGTATT CGGGGATGTT CACCTGGGAC AGAGTTGCAG ATTGAAGAGT   600
TCTACATCTC ACATCCTGTC ACACTATGTG TGATATTTAA GAAACTTTGT   650
TTGGCGGAAC AACAAGTTTG CACAAACATT TGAAGAAGAA AGCGAAATGA   700
TTCAGAGAG                                                709
```

Fig 1